**PREM PELLURU**

[**sprem656@gmail.com**](mailto:sprem656@gmail.com) **Contact no: (763) 392-2515**

**Hadoop Developer**

**Professional Summary:**

* Around **8+ years** of experience in all phases of SDLC including application design, development, production support & maintenance projects and Java with 3 years of experience in Hadoop eco system.
* Excellent understanding / knowledge of Hadoop architecture and various components such as HDFS, JobTracker, Task Tracker, NameNode, Data Node and MapReduce programming paradigm.
* Implemented Hadoop based data warehouses, integrated Hadoop with Enterprise Data Warehouse systems.
* Hands on experience in installing, configuring CDH 3 and CDH 4 clusters and using Hadoop ecosystem components like Hadoop MapReduce, HDFS, Pig, Hive, HBase, Cassandra, Sqoop, Oozie, Sentry, Flume, Zookeeper, Cloudera Hadoop Ecosystem.
* Experience on Apache Hadoop Map Reduce programming, PIG Scripting and Distribute Application and HDFS and in Apache Sentry for security applications in financial organizations.
* Good understanding of NoSQL Data bases and hands on work experience in writing applications onNoSQL databases like Cassandra and Mongo DB.
* Experience in managing and reviewing Hadoop log files.
* Experience in analyzing data using CQL, HiveQL, Pig Latin, HBase and custom Map Reduce programs in Java.
* Extending Hive and Pig core functionality by writing custom UDFs.
* Created UDFs to implement functionality not available in Pig. Used UDFs from Piggybank UDF Repository.
* Used the Hadoop MapReduce API to write programs in Java to perform custom JOINS, aggregates and groups.
* Experience in importing and exporting data using Sqoop from HDFS to Relational Database Systems and vice-versa.
* Hands on experience in application development using Java, RDBMS, and Knowledge in Linux shell scripting.
* Extensively worked on database applications using DB2, Oracle11g/10g, SQL, MySQL, PL/SQL, SQL\*PLUS.
* Strong experience as a senior Java Developer in Web/intranet, client/server technologies using Java, J2EE, Servlets, JSP, JSF, EJB, JDBC and SQL.
* Excellent working knowledge of popular frameworks like Struts, Hibernate, and Spring MVC
* Experience in Object Oriented Analysis, Design (OOAD) and development of software using UML Methodology, good knowledge of J2EE design patterns and Core Java design patterns.
* Solid understanding of the high volume, high performance systems
* Experience in managing and reviewing Hadoop log files.
* Excellent communication skills, interpersonal skills, problem solving skills, and a team player.

Technical Summary:

|  |  |
| --- | --- |
| Big Data Ecosystem | Apache HDFS, Apache Spark, Apache Sentry, Avro, Mahout MapReduce,HBase  Zookeeper, Hive, Pig, Sqoop, Oozie, Flume, Cassandra, CQL,CDH3, CDH4, Apache  Whirr, Apache Big Top, Apache Solr, Apache Nutch, Apache Lucene, Apache Giraph,  Apache Sentry, Apache Spark, Talend,. |
| Hadoop Distributions | Cloudera CDH 3, CDH 4, Amazon Web Services, Amazon EC2, S3 |
| Languages | C/C++, Core Java, SQL, PL/SQL, SQL |
| Methodologies | Agile, UML, Design Patterns |
| Database | Oracle 11g/10g, DB2, MySQL |
| Application Server | Apache Tomcat 5.x 6.0, Jboss 4.0 |
| Web Technologies | JDBC, JSP, JSF, EJB 2.0, XML, DTD,SOAP, Schemas, XSL,  XSLT, XPath, HTML |
| Frameworks | Struts, Spring |
| Analysis and Reporting  Tools | Microsoft SSRS2008/2012, Microsoft SSAS 2012,Splunk, Tableau, Pentaho,  Datameer |
| Tools | SQL developer, DB visualize |
| IDE / Testing Tools | NetBeans, Eclipse |
| Operating System | Windows, Linux Ubuntu 14.04, Cent OS 6.2 |
| Scripting Languages | Bash, Perl, Python, Ruby, JavaScript |
| Testing API | JUNIT |

**Professional Experience:**

1. **United Health Group, Plymouth, MN Hadoop Developer, BDIT March 2013-Till Date**

United Health Group, Inc is a diversified health care company that provides services across the nation on two different platforms i.e., UnitedHealthCare and Optum. Currently, UHG working with huge volume of claims data of almost a Petabyte range. In this project we mainly deal with claims data which is received from various providers and each member is linked with primary care physician using HQL. Based on Standard codes available for different diseases, we try to identify whether the member has corresponding problems or not using Pig and Hive. UHG sends multiple checks to each and every claim and we are asked to aggregate the check amount and send it as a single check.

**Responsibilities:**

* Gathered the business requirements from the Business Partners and Subject Matter Experts.
* Involved in installing Hadoop Ecosystem components.
* Involved with ingesting data received from various providers, on HDFS for big data operations.
* Loaded and transformed large sets of structured, semi structured and unstructured data in various formats like text, zip, XML and JSON.
* Wrote **MapReduce** jobs to perform big data analytics on ingested data using Java API.
* Supported MapReduce Programs those are running on the cluster.
* Installed and configured Pig and also written **Pig Latin** scripts.
* Imported data using **Sqoop** to load data from Oracle to HDFS on regular basisor from Oracle server to Hbasedepending on requirements.
* Developed Scripts and Batch Job to schedule various Hadoop Programs.
* Implemented advanced procedures like text analytics and processing using the in-memory computing capabilities like **Spark.**
* Enhanced and optimized product **Spark** code to aggregate, group and run data mining tasks using the Spark framework.
* Wrote MapReduce in **Ruby** using Hadoop Streaming to implement various functionalities.
* Wrote Hive queries for data analysis to meet the business requirements.
* Created Hive tables and working on them using **Hive QL**.
* Loaded the data into **Cassandra**using **CQL** and carefully, generated the rowkey.
* Built web portal using JavaScript, it makes a **REST API** call to the elastic search and gets the rowkey.
* Built Hive table on top of Cassandra so that it can point to the reporting tools like **Tableau, Datameer, Splunk**.
* Used **netezza** for low latency queries by loading the data into netezza from bigdata cluster using**nzload**.
* Used **Kibana**, which is an open source based browser analytics and search dashboard for Elastic Search.
* Utilized Agile Scrum Methodology to help manage and organize a team of 4 developers with regular code review sessions.
* Weekly meetings with technical collaborators and active participation in code review sessions with senior and junior developers.

**Environment**: **Cassandra**, **netezza**, Pig, Hive, MapReduce, **Sqoop**, Ruby, JavaScript, Apache Spark,Tableau, Talend, Elastic Search, Kibana, Datameer, Splunk, Pentaho,SFTP.

1. **Bank of America, Charlotte, NC Hadoop Developer March 2012 – February 2013**

### **Operational Data Store** is a hub for all the tables that are used in Bank, Card and Mortgage Applications. There are 58 applications comes under ODS. This Project deals with developing historical database using Hadoop ecosystem for maintaining last 10 years of data spread across branches in US. Main aim of the project is to centralize the source of data for audit/legal report generation using historical database which otherwise are generated from multiple sources.

**Responsibilities:**

* Worked on Hadoop cluster which ranged from 8-10 nodes during pre-production stage and it was sometimes extended up to 40 nodes during production
* Installed and configured Hadoop **MapReduce**, HDFS, Developed multiple **MapReduce**jobs in java for data cleaning and preprocessing.
* Established custom MapReduce programs in order to analyze data and used **Pig Latin** to clean unwanted data
* Did various performance optimizations like using distributed cache for small datasets, Partition, Bucketing in hive and Map Side joins.
* Built reusable **Hive** UDF libraries.
* Used**Sqoop** to import the data from RDBMS to Hadoop Distributed File System (HDFS) and later analyzed the imported data using Hadoop Components.
* Created Hive tables, then applied **HiveQL** on those tables, this will invoke and run **MapReduce** jobs automatically
* Performed loading and transforming large sets of Structured, Semi-Structured and Unstructured data and analyzed them by running Hive queries and Pig scripts
* Installed **Oozie workflow** engine to run multiple Hive and Pig jobs which run independently with time and data availability
* Performed Hadoop installation, updates, patches and version upgrades when required
* Performed Cluster maintenance, Cluster Monitoring and Troubleshooting, Manage and review data backups and log files
* Bug fixing and 24x7production support

**Environment:** CDH3,CDH4, PIG(0.8.1), HIVE(0.7.1), **Sqoop** (V1),**Java**, **Eclipse**, Flume, Zookeeper, Oozie, Oracle 11g, SQL Server 2008, **HBase**, Oracle 11g / 10g,PL/SQL, SQL\*PLUS, Toad 9.6, Windows NT, LINUX, UNIX Shell Scripting.

1. **Intuit Inc., San Diego, CA Java/Hadoop Developer November2010 – Feb 2012**

Intuit Inc., is an American Software company that develops financial and tax preparation software and related services for small businesses, accountants and individuals which helps in managing their personal finances quickly and easily and able to manage the thing as real bank. The **Mint**application manages multiple accounts of different types like checking, savings, credit cards, loans, debt accounts. The application has the facility to create own account type, like cash allowance for even small Activities. The Applicationhas the Flexibility to Setup budget limits for every category and track the difference between actual and budget values. **Intuit** also provides **Turbo Tax, Quicken, Quicken Books** which helps in filing the tax forms for individuals and small scale industries with great refunds. This project mainly deals with users of Mint and Turbo Tax software applications with more than 10 million users whose information is all set up in a secured way in Cloudera CDH4 cluster and also which put all these information in a secured way using **Apache Sentry.**

**Responsibilities**:

* Launching and Setup of HADOOP Cluster which ranged 40 to 60 nodes on Cloudera CDH 4, which includes configuring different components of HADOOP.
* Experience in Using **Sqoop** to connect to the **DB2** and move the pivoted data to **Hive** tables or **Avro** files.
* Managed the Hive database, which involves ingest and index of data.
* Expertise in exporting the data from Avro files and indexing the documents in sequence or **SerDe** file format.
* Hands on experience on **Oozie** workflow.
* Responsible for Analysis, Design, Development and Integration of UI components with backend usingJ2EE technologies such as **Servlets, Java Beans, JSP, JDBC**.
* Used **Spring Framework 3.2.2** for transaction management.
* Used **Spring**for Transaction management and **Hibernate3** to persist the data into the database.
* Writing JSP's for user interfaces, **JSP**'s uses Java Beans objects to produce responses.
* Created controller Servlets for handling **HTTP** requests from JSP pages.
* Writing JavaScript functions for various validation purposes. Writing Java Beans to contain data objects.
* Implemented the presentation layer using **Struts2 MVC** framework.

**Environment:** Hive, Sqoop, Oozie, HDFS, MapReduce, Bit bucket, Maven, J2EE, Struts, Spring, Unix, SQL,

JSP, Servlets, JDBC, Hibernate, Java Beans.

1. **Infosys, Bangalore, India Sr. Java DeveloperMarch 2009- October 2010**

This project intended at building a web based Billing and Customer Care System. Customers can work with subsystems, such as Registration, current policy information, getting a quote, make changes to his existing policy, apply for a new policy, make payment, checking for claims status, place a claim and others.

**Responsibilities:**

* Developed all the User Interfaces using **JSP** and **Struts** framework.
* Writing Client Side validations using JavaScript.
* Extensively used JQuery for developing interactive web pages.
* Developed the DAO layer using **Hibernate** and for real time performance used the caching system for hibernate.
* Experience in developing web services for production systems using SOAP and WSDL.
* Developed the user interface presentation screens using HTML, XML, and CSS.
* Experience in working with **spring** using **AOP**, **IOC** and **JDBC** template.
* Developed the Shell scripts to trigger the Java **Batch** job, Sending summary email for the batch job status and processing summary.
* Co-ordinate with the QA lead for development of test plan, test cases, test code and actual testing responsible for defects allocation and those defects are resolved.
* Involved in testing and deployment of the application on Web logic Application Server during integration and QA testing phase.
* Maintained the existing code base developed in the **Struts**, **spring** and **Hibernate** framework by incorporating new features and doing bug fixes.
* Used ANT tool to build and deploy applications.
* Involved in configuring web.xml and struts.xml for workflow.
* Wrote SQL queries and created DDL scripts for interacting with the Oracle database.
* Was part of solving the issues that came in production.
* Documentation of common problems prior to go-live and while actively in a Production Support role.

**Environment:** Java, J2EE, Spring, Struts, J2SE, JSP, SOAP, EJB, WDSL, Hibernate, Servlet, HTML, XML, CSS, JavaScript, Eclipse, Oracle, Tomcat, Linux, TCP/IP.

1. **HDFC Bank, Bangalore, India Java Developer October 2007 – February 2009**

Wide arena of financial products and services, HDFC Group has solutions like InstaBanking, Online Trading, Insta Insure, HDFC Bank immobile etc. Providing high class financial services in all segments of the society, HDFC Group deals with Mutual Fund, Private Equity, Securities and Life Insurance, etc. This project mainly deals with Credit Cards security features and Consumer Banking.

**Responsibilities**:

* Involved in application architecture design phase and creatingthe application design template.
* Responsible to get the requirements from Onsite coordinator /customer.
* Responsible to prepare the Low Level design document andcreating the reusable components for across the application
* Assisted with code reviews,conducted technology discussion groups and presentations, act asresource for them.
* Prepared detailed design documents of each screen.
* Responsibility to do the Code Review and preparing the IQA forother application components.
* Participating in customer meeting to identifying thedifficulties if any, and discuss about the application status and get the requirements.
* Responsible for Detail design and GUI development.
* Worked on **Spring** and **Hibernate** mapping using annotation.
* Utilized Java and MySQL from day to day to debug and fix issues with client processes.
* Developed, tested, and implemented financial‐services application to bring multiple clients intostandard database format.
* Assisted in designing, building, and maintaining database to analyze life cycle of checking anddebit transactions.
* Excellent JAVA, J2EE application development skills with strong experience in Object OrientedAnalysis.
* Extensively involved throughout Software Development Life Cycle (SDLC).
* Strong experience of **J2SE, XML, Web Services, SOAP, TCP, IP**.
* Strong experience of software and system development using **JSP, Servlet, Java Server Face,EJB, JDBC, Struts, Subversion, JUnit, SQL language.**
* Rich experience of database design and hands‐on experience of large database systems: **Oracle8i and Oracle 9i, DB2, PL/SQL.**
* Hands-on experience of Sun One Application Server, Web logic Application Server, Web Sphere
* Application Server, Web Sphere Portal Server, and J2EE application deployment technology.
* Strong problem solving, analytical and Communication skills.

**Environment**: Java, J2EE, Spring, Struts,Junit, J2SE, SOAP, EJB, JDBC,Hibernate, Servlet, XML, Eclipse, Oracle, DB2 Tomcat, Linux, TCP/IP.

1. **Mind Tree Ltd., Bangalore, India Java Developer June 2006 – Sep 2007**

Mind tree Ltd delivers technology services and accelerates growth for Global 1000 companies by solving complex business challenges with breakthrough technical innovations. Mind tree specializes in e-commerce, mobility, cloud enablement, digital transformation, business intelligence, data analytics, testing, infrastructure, EAI and ERP solutions.

**Responsibilities:**

* Developed front-end screens using **JSP**, **HTML** and **CSS**.
* Developed server side code using **Struts** and **Servlets**.
* Developed core java classes for exceptions, utility classes, business delegate, and test cases.
* Developed SQL queries using **MySQL** and established connectivity using **Hibernate.**
* Worked with Eclipse using **Maven** plugin for **Eclipse IDE**.
* The application was developed in Eclipse IDE and was deployed on **Tomcat server**.
* Involved in scrum methodology.
* Supported for bug fixes and functionality change.

**Environment**:Java Struts 1.1, Servlets, JSP, HTML, CSS, JavaScript, Eclipse 3.2, Tomcat, Maven 2.x, MySQL, Windows, Linux.

Education:

Bachelor of Engineering in Electronics and Communication, Anna University, Chennai, India

Certifications:

IBM **BigDataUniversity** certification in **Hadoop fundamentals, MapReduce programming, Hive, Pig, Oozie, Sqoop, Flume, Hbase**